

Index

- ℓ_1 norm, 59
- ℓ_1 regularization, 62
- ℓ_2 norm, 60
- ℓ_2 regularization, 62
- ℓ_∞ norm, 60
- ℓ_p norm, 60
- ε -greedy, 194
- action, 193
- action value function, 207
- activation units, 9
- actor, 239
- actor–critic methods, 239
- AdaDelta, 47
- adaptive gradient descent, 45
- adaptive moment estimation (Adam), 47
- adaptive subgradient descent (Adagrad), 47
- advantage actor–critic (A2C), 240
- agent, 193
- agent action, 193
- AlphaZero, 245
- analytic gradient, 38
- asynchronous advantage actor–critic (A3C), 240
- attention, 118
- auto-regressive Transformers, 147
- autoencoder, 181
- autoencoding Transformers, 146
- backpropagation, 21
- backpropagation through time, 99
- backtrack line search, 43
- bag of words, 91
- batch normalization, 65
- Bellman expectation equation, 213
- Bellman optimality equation, 215
- best step-size optimization, 43
- BFGS correction, 54
- bias, 59
- bidirectional encoder representations from Transformers (BERT), 143
- bidirectional recurrent neural network, 98
- bottleneck, 181
- Bregman divergence, 157, 176
- chain rule for differentiation, 26, 30
- co-evolution, 155
- conditional GAN (CGAN), 164
- contrastive language-image pre-training (CLIP), 269
- convex optimization, 35
- convolution, 70
- convolution kernel, 70
- convolution layer, 78
- convolution padding, 70
- convolution stride, 73
- convolutional neural network (CNN), 80
- covariance matrix adaptation, 55
- critic, 239
- cross validation, 59
- cycle consistent GAN (CycleGAN), 165
- data augmentation, 56, 64
- deep Q -network (DQN), 233
- deep convolutional GAN (DCGAN), 164
- deep deterministic policy gradient (DDPG), 244
- denoising diffusion probabilistic model (DDPM), 186
- DenseNet, 83
- derivative, 37
- derivative of hyperbolic tangent function, 27
- derivative of log-likelihood function, 178
- derivative of ReLU, 27
- derivative of sigmoid function, 26
- deterministic policy, 203
- DFP correction, 54
- discount factor, 199, 209
- double DQN, 235
- dropout, 56, 62
- dueling network, 235
- dynamic programming, 218
- Earth mover’s distance (EMD), 162
- eligibility traces, 227
- environment, 193
- environment model, 210
- episode, 211
- evidence lower bound (ELBO), 177
- evolution strategies, 37, 54
- expected log likelihood, 178

- expected return, 212
- expected reward, 193
- experience replay, 232
- expert iteration, 246
- exploitation, 211
- exploration, 211
- exponential family of distributions, 175
- exponentially weighted moving average, 46
- f-divergence, 175
- f-GAN, 158
- farsighted agent, 206
- feature vector sentence representation, 92
- finite horizon, 204
- first-order methods, 37
- Fisher information matrix, 242
- forward propagation, 11
- Frechet inception distance (FID), 168
- fully connected neural network, 9
- GAN discriminator, 156
- GAN discriminator training, 160
- GAN discriminator–generator training, 161
- GAN generator, 156
- GAN generator training, 160
- GAN minimax loss, 162
- GAN training, 160
- gated graph neural networks, 139
- gated recurrent unit (GRU), 102
- generalization gap, 56
- generative adversarial network (GAN), 153
- Go-Explore, 249
- GPT-3, 148
- gradient, 38
- gradient descent, 32, 36, 39
- gradient descent ascent (GDA), 158
- gradient descent of loss function, 33
- gradient descent update, 33
- gradient descent with momentum, 46
- gradient of ELBO, 178
- graph, 126
- graph adjacency list, 126
- graph adjacency matrix, 126
- graph attention network (GAT), 140
- graph connected components, 129
- graph convolution network (GCN), 138
- graph edge, 126
- graph Laplacian matrix, 127
- graph neighborhood aggregation, 138
- graph neighborhood aggregation function, 139
- graph neural network (GNN), 136
- graph node, 126
- graph node degree, 127
- graph node embedding, 130
- graph node feature vector, 131
- graph node similarity, 131
- graph random walk, 133
- graph random walk normalized Laplacian matrix, 128
- graph shallow node embedding, 131
- graph symmetric normalized Laplacian matrix, 128
- graph walk, 129
- graphSAGE, 139
- greedy action, 194
- Hessian, 38, 53
- history, 211
- huggingface Transformers platform, 150
- hyperbolic tangent function, 16
- hypergradient descent, 48
- image-to-image translation (Pix2Pix), 165
- imitation learning, 248
- inception score (IS), 168
- infinite horizon, 206
- initial state, 196
- input normalization, 33
- instance conditioned GAN (IC-GAN), 168
- inverse document frequency, 91
- inverse Hessian, 53
- inverse reinforcement learning, 248
- invertible residual neural network, 86
- iterative policy evaluation, 218
- Jacobian matrix, 31
- Jensen–Shannon (JS) divergence, 157
- joint density, 174
- Kullback–Leibler (KL) divergence, 157, 176
- L-BFGS, 54
- Lasso, 61
- latent representation, 183
- latent variable, 174
- learning curves, 58
- learning rate, 41
- least-squares GAN (LS-GAN), 158
- likelihood function, 174
- local maximum, 38
- local minimum, 38
- logistic regression, 14
- logistic regression loss, 21
- long short-term memory (LSTM), 108
- loss function, 19, 35
- marginal density, 174
- Markov decision process (MDP), 199
- Markov model, 92
- Markov process, 197
- max pooling, 78
- mean squared error, 20
- message-passing graph neural network, 140
- mini-batch gradient descent, 44
- minimax optimization problem, 155
- MobileNet, 85
- model-based reinforcement learning, 210, 245
- model-free reinforcement learning, 211, 232

- Monte Carlo sampling, 222
- multi-armed bandit, 193
- multi-head attention, 144
- multi-modal Transformer, 149
- myopic agent, 206
- n-gram, 92
- negative entropy, 178
- Nesterov momentum, 46
- neural fitted Q -iteration, 233
- neural network, 9
- neural network layers, 9
- Newton's method, 49
- Newton–Raphson update, 52
- non-convex optimization, 36
- non-linear activation function, 16
- normalizing flows, 184
- numerical gradient, 38
- observation, 196
- ODENet, 85
- off-policy reinforcement learning, 227
- on-policy reinforcement learning, 227
- OpenAI Codex, 149
- optimal action value function, 215
- optimal infinite horizon policy, 215
- optimal policy, 214
- optimal state value function, 215
- optimistic gradient descent ascent (OGDA), 159
- optimization problem, 35
- overfitting, 56
- planning, 218
- policy, 198, 202
- policy evaluation, 218
- policy gradient, 237
- policy improvement, 218
- policy iteration, 218
- policy search, 222
- policy-based reinforcement learning, 236
- pooling, 78
- posterior, 174
- pre-activations, 9
- proximal policy optimization (PPO), 244
- PyTorch, 33
- Q -learning, 224
- quasi-Newton methods, 36, 53
- rectified linear unit (ReLU), 17
- recurrent hidden units, 94
- recurrent loss function, 95
- recurrent neural network (RNN), 93
- registration GAN (RegGAN), 167
- regularization, 56
- regularization parameter, 61
- regularized loss functions, 61
- REINFORCE algorithm, 238
- reinforcement learning, 193
- reparameterization gradient, 180
- replay buffer, 233
- residual neural network (ResNet), 81
- return, 210
- reward, 193, 209
- reward matrix, 200
- ridge regression, 60
- Riemannian manifold, 187
- RMSProp, 47
- saddle-point, 38
- saddle-point problem, 155
- score function, 179
- secant condition, 53
- second derivative, 37
- second-order methods, 37, 49
- self-attention, 144
- self-attention GAN (SAGAN), 167
- semi-supervised GAN (SGAN), 164
- sequence-to-sequence, 117
- sequence-to-sequence Transformers, 147
- sigmoid function, 16
- softmax function, 18
- squeeze and excitation network (SENet), 84
- SR1 update, 54
- state, 200
- state machine, 92, 196
- state of environment, 193, 201
- state value function, 204
- state–action diagram, 203
- step size, 41
- stochastic gradient descent, 28, 44
- stochastic policy, 203
- stochastic weight averaging, 45
- sub-graph embedding, 131
- Swish function, 18
- target network, 234
- temporal difference (TD) learning, 222
- TensorFlow, 33
- term frequency, 91
- test accuracy, 58
- TFIDF, 92
- training data, 58
- training neural networks, 45
- Transformer, 142
- Transformer decoder, 146
- Transformer encoder, 145
- Transformer fine-tuning, 146
- Transformer masked word prediction, 146
- Transformer position encoding, 145
- Transformer pre-training, 146
- transition function, 196
- transition matrix, 198
- transition model, 199
- trust region policy optimization (TRPO), 243
- upper confidence bound (UCB), 196

value function approximation, 230
value iteration, 219
variance, 59
variational autoencoder (VAE), 183
variational distribution, 175
variational inference (VI), 175
variational inference problem, 175
vector norms, 59
vision Transformer (ViT), 148
Wasserstein GAN (WGAN), 162
Wasserstein GAN loss, 162
Wasserstein-1 distance, 162
weight initialization, 33
word embedding, 121
world models, 245